بهینه‌سازی (۲)

در این جلسه نخست به تشریح نقش فرض $\nabla g(a) \neq \underline{\circ}$ در قضیهٔ لاگرانژ می‌پردازیم. همان طور که مثال آخر جلسهٔ قبل نشان داد وقتی $\nabla g(a) = \underline{\circ}$، ممکن است مجموعهٔ تراز تابع g که از a می‌گذرد در نقطهٔ a هموار نباشد و در این صورت مماس و قائم در آن نقطه فاقد معنی می‌شوند. یادآوری می‌کنیم که "هموار بودن" مجموعه تراز $M = g^c$ در نقطهٔ a را بدین معنی گرفتیم که بتوانیم بخشی از مجموعهٔ M حول a را به صورت نمودار یک تابع مشتق‌پذیر $(n-۱)$ متغیری بنویسیم. حال $\nabla g(a) = (\frac{\partial g}{\partial x_۱}(a), \ldots, \frac{\partial g}{\partial x_n}(a))$ و شرط $\nabla g(a) \neq \circ$ معادل این است که دست کم یکی از $\frac{\partial g}{\partial x_i}(a)$ ها صفر نباشد. مثلاً فرض کنید $\frac{\partial g}{\partial x_n}(a) \neq \circ$. در این صورت بنابر قضیهٔ تابع ضمنی نقاط $(x_۱, \ldots, x_n)$ از مجموعهٔ تراز $M = g^c$ را که نزدیک a هستند می‌توان به صورت نمودار تابعی مشتق‌پذیر $x_n = \varphi(x_۱, \ldots, x_{n-۱})$ نمایش داد و تحدید تابع f به M را می‌توان به صورت تابعی از $x_۱, \ldots, x_{n-۱}$ نوشت:

$$f(x_۱, \ldots, x_n) = f(x_۱, \ldots, x_{n-۱}, \varphi(x_۱, \ldots, x_{n-۱}))$$

که عبارت سمت راست را به $h(x_۱, \ldots, x_{n-۱})$ نمایش می‌دهیم. حال اگر a یک نقطهٔ ماکسیمم یا مینیمم موضعی تحدید f به M باشد، $a' = (a_۱, \ldots, a_{n-۱})$ یک نقطهٔ بحرانی برای h است، و در واقع در هر نقطهٔ بحرانی a' از h داریم $\nabla h(a') = \underline{\circ}$، یا $\frac{\partial h}{\partial x_i}(a') = \circ$ برای هر i. طبق قاعدهٔ زنجیره‌ای داریم:

$$\frac{\partial h}{\partial x_i} = \sum_{j=۱}^{n-۱} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial \varphi}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial \varphi}{\partial x_i}$$

پس در نقطهٔ بحرانی a:

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial \varphi}{\partial x_i} = \circ \quad , \quad i = ۱, \ldots, n-۱ \qquad (۱)$$

از طرفی دیگر چون مجموعهٔ M در واقع مجموعهٔ تراز منسوب به 0 تابع $\varphi(x_1,\dots,x_{n-1})-x_n$ است، جهت عمود بر آن توسط گرادیان این تابع، یعنی بردار زیر داده می‌شود:

$$(\frac{\partial\varphi}{\partial x_1},\dots,\frac{\partial\varphi}{\partial x_{n-1}},-1) \tag{2}$$

از مقایسهٔ روابط (۱) و (۲) نتیجه می‌شود که:

$$(\frac{\partial\varphi}{\partial x_1}:\dots:\frac{\partial\varphi}{\partial x_{n-1}}:-1)=(\frac{\partial f}{\partial x_1}:\dots:\frac{\partial f}{\partial x_{n-1}}:\frac{\partial f}{\partial x_n})$$

پس در نقطهٔ a، $\bigtriangledown f(a)$ عمود بر M است. استدلال بالا در واقع اثبات دیگری از قضیهٔ لاگرانژ است که در آن به صراحت نشان داده شد که هرگاه مجموعهٔ تراز M در a هموار باشد و a یک نقطهٔ بحرانی برای تحدید f به M، آنگاه $\bigtriangledown f(a)$ بر مجموعهٔ تراز M عمود است.

هدف بعدی ما تعمیم قضیهٔ لاگرانژ به حالتی است که به جای $g(x_1,\dots,x_n)=c$، تعدادی رابطه از این نوع ملحوظ شده باشد. به طور کلی، به خصوص در کاربردها، اصطلاحات زیر معمول است. تابع f که ماکسیمم یا مینیمم تحدید آن به زیرمجموعهٔ خاصی مطرح است <u>تابع عینی</u>، و شرط $g(x_1,\dots,x_n)=c$ (تعریف کنندهٔ مجموعهٔ تراز $M=g^c$) را یک <u>قید</u> می‌نامند. تا اینجا فقط مسائل تک قیدی را مطرح کرده‌ایم و اکنون قصد داریم مسائل چند قیدی را نیز بررسی کنیم.

وضعیت زیر را در نظر می‌گیریم. تابع $f:U\longrightarrow\mathbb{R}$ داده شده است که تابعی مشتق‌پذیر روی زیرمجموعهٔ باز U از $\mathbb{R}^n$ است. به علاوه تابع‌های $g_1,\dots,g_k:U\longrightarrow\mathbb{R}$ نیز داده شده‌اند که دارای مشتق‌های پاره‌ای مرتبهٔ اول پیوسته می‌باشند. می‌خواهیم ماکسیمم یا مینیمم تحدید تابع f به اشتراک مجموعه‌های تراز $M_1:g_1(x_1,\dots,x_n)=c_1$، ...، $M_k:g_k(x_1,\dots,x_n)=c_k$ را بیابیم. به بیان دیگر، مقصود یافتن ماکسیمم یا مینیمم f تحت قیود $g_1(x_1,\dots,x_n)=c_1$، ...، $g_k(x_1,\dots,x_n)=c_k$ است. در حالت $k=1$ دیدیم که اگر در نقطهٔ مطلوب a، $\bigtriangledown g(a)\neq 0$، آنگاه مجموعهٔ تراز $M:g(x)=c$ حول نقطهٔ a هموار است، و $\bigtriangledown f(a)$ عمود بر این مجموعه عمود خواهد بود. حال اگر اشتراک مجموعه‌های تراز حول a هموار باشد، آنگاه خواهیم دید که به دلیلی کاملاً مشابه حالت تک قیدی، $\bigtriangledown f(a)$ بر این اشتراک عمود است. بدین ترتیب دو سؤال زیر مطرح می‌شود:

۱) تحت چه شرایطی اشتراک مجموعه‌های تراز حول a هموار است؟

۲) به شرط هموار بودن اشتراک مجموعه‌های تراز حول a، چگونه می‌توان عمود بودن $\nabla f(a)$ بر این اشتراک را به صورت یک رابطهٔ جبری قابل استفاده بیان کرد؟

نخست سؤال (۱) را بررسی می‌کنیم. اشتراک مجموعه‌های تراز، مجموعهٔ $(x_1, \ldots, x_n)$ هایی است که در روابط زیر صدق می‌کنند:

$$\begin{cases} g_1(x_1, \ldots, x_n) - c_1 = 0 \\ \vdots \\ g_k(x_1, \ldots, x_n) - c_k = 0 \end{cases}$$

فرض کنید نقطهٔ a در این روابط صدق کند. اگر بتوان از این روابط، حول a، k متغیر را به صورت تابع‌های مشتق‌پذیر از $(n-k)$ متغیر باقیمانده نوشت، آنگاه بخشی از اشتراک $M_1 \cap \ldots \cap M_k$ که حول a قرار دارد به صورت نمودار یک تابع مشتق‌پذیر $(n-k)$ متغیری ظاهر می‌شود و بدین معنی "هموار" است. طبق قضیهٔ تابع ضمنی اگر یک زیرماتریس $k \times k$ از ماتریس $k \times n$ زیر، متشکل از k ستون، دارای دترمینان ناصفر باشد، آنگاه k متغیر مربوط به این k ستون را می‌توان به صورت تابع‌های مشتق‌پذیر نسبت به $(n-k)$ متغیر باقیمانده نوشت:

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1}(a) & \cdots & \frac{\partial g_k}{\partial x_n}(a) \end{bmatrix} \qquad (۳)$$

توجه کنید که ردیف iـ ام این ماتریس از مؤلفه‌های $\nabla g_i(a)$ تشکیل شده است. در حالت $k = 1$، که قبلاً بررسی شد، وجود یک ماتریس 1×1 با دترمینان مخالف صفر در واقع معادل این است که $\nabla g(a)$ صفر نباشد (یعنی حداقل یک مؤلفهٔ ناصفر داشته باشد). برای $k > 1$ توجه کنید اگر یکی از $\nabla g_i(a)$ ها را بتوان به صورت "ترکیبی خطی" از سایر $\nabla g_i(a)$ ها نوشت، یعنی مثلاً:

$$\nabla g_k(a) = c_1 \nabla g_1(a) + \cdots + c_{k-1} \nabla g_{k-1}(a)$$

که در آن $c_1, \ldots, c_{k-1}$ عدد حقیقی هستند، آنگاه برای هر زیرماتریس $k \times k$ از ماتریس بالا نیز، یک ردیف ترکیبی خطی از سایر ردیف‌ها می‌شود و در نتیجه دترمینان صفر خواهد شد. طبق قضایای جبر خطی، عکس موضوع فوق نیز صحیح است، یعنی شرطی لازم و کافی برای این که ماتریس (۳) دارای

زیر ماتریس $k \times k$ با دترمینان ناصفر باشد این است که نتوان هیچ‌یک از $\nabla g_i(a)$ ها را به صورت ترکیبی خطی از سایر $\nabla g_i(a)$ ها نوشت.

بدین ترتیب جایگزین مناسب برای شرط $\nabla g(a) \neq 0$ در حالتی که k قید موجود باشند این خواهد بود که $\{\nabla g_1(a), \ldots, \nabla g_k(a)\}$ مستقل خطی باشد. این شرط تضمین می‌کند که بخشی از $M = M_1 \cap \ldots \cap M_k$ که در نزدیکی a قرار دارد نمودار یک تابع $(n-k)$ متغیری مشتق‌پذیر است و به این مفهوم از "هموار بودن" برخوردار است.

حال به سؤال دوم می‌پردازیم. فرض کنید اشتراک $M = M_1 \cap \ldots \cap M_k$ حول a هموار است و بردار w بر این اشتراک عمود می‌باشد. با توجه به ملاحظات بالا که M حول a نمودار یک تابع مشتق‌پذیر $(n-k)$ـ متغیری است و با $(n-k)$ پارامتر به طور مشتق‌پذیر توصیف می‌شود، می‌توان آن را یک شیء هموار $(n-k)$ـ بعدی تصور کرد که در $\mathbb{R}^n$ قرار گرفته است. بنابراین جهت عمود بر M در نقطهٔ a، k بعدی می‌شود. توجه کنید که k بردار مستقل خطی $\nabla g_1(a), \ldots, \nabla g_k(a)$ همه در راستای عمود بر M قرار دارند زیرا که $\nabla g_i(a)$ بر M_i عمود است و در نتیجه هر زیرمجموعه M از M_i نیز عمود می‌باشد. پس ترکیب‌های خطی $c_1 \nabla g_1(a) + \cdots + c_k \nabla g_k(a)$ یک فضای kـ بعدی پدید می آورند که همان فضای kـ بعدی عمود بر M است. بدین ترتیب باید بتوان w را به صورت ترکیبی خطی از $\nabla g_1(a), \ldots, \nabla g_k(a)$ نوشت. با توجه به این توضیحات، اکنون آماده‌ایم تعمیم قضیهٔ لاگرانژ را ارائه کنیم:

(۱۷–۱)<u>قضیه لاگرانژ (صورت کامل)</u> U زیرمجموعه‌ای باز از $\mathbb{R}^n$ است، $f : U \longrightarrow \mathbb{R}$ تابعی مشتق‌پذیر، و $g_1, \ldots, g_k : U \longrightarrow \mathbb{R}$ تابع‌هایی با مشتق‌های پاره‌ای مرتبهٔ اول پیوسته. فرض کنید برای اعداد حقیقی داده شدهٔ $c_1, \ldots, c_k$، مجموعه‌های $M_1, \ldots, M_k$، به ترتیب، مجموعه‌های تراز $g_1, \ldots, g_k$ منسوب به $c_1, \ldots, c_k$ باشند، یعنی $M_i = g_i^{c_i}$، و $M = M_1 \cap \ldots \cap M_k$. اگر نقطهٔ a در M یک نقطهٔ ماکسیمم یا مینیمم موضعی تحدید تابع f به M باشد و اگر $\nabla g_1(a), \ldots, \nabla g_k(a)$ مستقل خطی باشند، آنگاه اعداد حقیقی $\lambda_1, \ldots, \lambda_k$ وجود دارند که:

$$\nabla f(a) = \lambda_1 \nabla g_1(a) + \cdots + \lambda_k \nabla g_k(a) \qquad (۴)$$

با توجه به مقدمات ذکر شده بیشتر اثبات این قضیه قبلاً آماده شده است، فقط باید ثابت کنیم که اگر a یک نقطهٔ ماکسیمم یا مینیمم موضعی برای تحدید f به M باشد، آنگاه $\nabla f(a)$ بر M عمود است. ولی اثبات این مطلب عیناً مانند حالت $k = ۱$ است. اگر $\gamma : I \longrightarrow \mathbb{R}^n$ خمی مشتق‌پذیر باشد که تصویر آن به تمامی در M قرار می‌گیرد و $a = \gamma(t_\circ)$، باید ثابت کنیم $\nabla f(a)$ بر $\gamma'(t_\circ)$ عمود است. برای این کار توجه می‌کنیم که $f \circ \gamma$ در $t_\circ$ دارای ماکسیمم یا مینیمم موضعی است، پس $(f \circ \gamma)'(t_\circ) = \circ$، و طبق قاعدهٔ زنجیره‌ای $\nabla f(a) \cdot \gamma'(t_\circ) = \circ$ و حکم به اثبات می‌رسد.

در زیر به ذکر دو مثال می‌پردازیم. توجه کنید که (۴) در واقع n معادله با $(n+k)$ مجهول $x_۱, \ldots, x_n, \lambda_۱, \ldots, \lambda_k$ به‌دست می‌دهد، و با کمک k قید داده شده مجموعاً $(n+k)$ معادله با $(n+k)$ مجهول در دست است.

<u>مثال ۱</u> برای اشتراک استوانهٔ $x^۲ + y^۲ = ۱$ و صفحهٔ $x + ۲y + ۳z = ۱$، نقطه‌هایی که بزرگترین و کوچکترین z را دارند پیدا کنید.

<u>حل</u> تابع عینی در اینجا $f(x,y,z) = z$ است و قیدها عبارتند از $x^۲ + y^۲ - ۱ = \circ$ و $x + ۲y + ۳z - ۱ = \circ$. می‌نویسیم $g_۱(x,y,z) = x^۲ + y^۲ - ۱$ و $g_۲(x,y,z) = x + ۲y + ۳z - ۱$. داریم $\nabla g_۱(x,y,z) = (۲x, ۲y, \circ)$ و $\nabla g_۲(x,y,z) = (۱,۲,۳)$. $\nabla g_۲$ که هیچ‌گاه مضربی از $\nabla g_۱$ نیست زیرا که مؤلفه سوم $\nabla g_۲$ برابر ۳ است و مؤلفه سوم $\nabla g_۱$ برابر ۰. تنها به‌ازای $x = y = \circ$، $\nabla g_۱$ مضرب (صفر) $\nabla g_۲$ می‌شود، ولی هیچ نقطهٔ به شکل $(\circ, \circ, z)$ روی استوانهٔ داده شده قرار ندارد، پس شرط استقلال خطی در همهٔ نقاط اشتراک برقرار است. بدین ترتیب می‌توانیم از رابطهٔ $\nabla f = \lambda_۱ \nabla g_۱ + \lambda_۲ \nabla g_۲$ استفاده کنیم:

$$\begin{cases} \circ = \lambda_۱(۲x) + \lambda_۲(۱) \\ \circ = \lambda_۱(۲y) + \lambda_۲(y) \\ ۱ = \lambda_۱(\circ) + \lambda_۲(۳) \end{cases}$$

از معادله سوم نتیجه می‌شود که $\lambda_۲ = \frac{۱}{۳}$ و با جایگزین در معادله اول معلوم می‌شود که $\lambda_۱ \neq \circ$، پس $x = \frac{-۱}{۶\lambda_۱}$ و $y = \frac{-۱}{۳\lambda_۱}$. با جایگزین در $x^۲ + y^۲ = ۱$ داریم $\lambda_۱ = \pm\frac{\sqrt{۵}}{۶}$، پس با استفاده

از $x + ۲y + ۳z = ۱$، دو نقطهٔ $(\frac{۱}{\sqrt{۵}}, \frac{۲}{\sqrt{۵}}, \frac{۱-\sqrt{۵}}{۳})$ و $(\frac{-۱}{\sqrt{۵}}, \frac{-۲}{\sqrt{۵}}, \frac{۱+\sqrt{۵}}{۳})$ به‌دست می‌آیند، که اولی بزرگترین z و دومی کوچکترین z ممکن را دارد.

مثال ۲ فرض کنید که $g_۱(x,y,z) = ۰$ و $g_۲(x,y,z) = ۰$ دو رویه هموار در $\mathbb{R}^۳$ هستند که اشتراک تهی دارند، و نقطه P روی رویهٔ اول و نقطهٔ Q روی رویهٔ دوم طوری هستند که پاره‌خط PQ کوتاهترین طول ممکن میان پاره‌خط‌های واصل بین دو رویه را دارد. نشان دهید پاره‌خط PQ بر هر دو رویه عمود است.

حل هموار بودن دو رویه را بدین معنی می‌گیریم که تابع‌های $g_۱$ و $g_۲$ دارای مشتق‌های پاره‌ای مرتبه اول پیوسته‌اند، و $\nabla g_۱$ در سراسر $g_۱(x,y,z) = ۰$ و $\nabla g_۲$ در سراسر $g_۲(x,y,z) = ۰$ ناصفر هستند. اگر یک نقطهٔ رویهٔ اول را به (x,y,z) و یک نقطهٔ رویهٔ دوم را به (x',y',z') نمایش دهیم، تابعی که مینیمم آن مطرح است تابع ٦ متغیری $f(x,y,z,x',y',z') = (x-x')^۲ + (y-y')^۲ + (z-z')^۲$ می‌باشد. تابع‌های $g_۱$ و $g_۲$ را به صورت زیر بازنویسی می‌کنیم:

$$G_۱(x,y,z,x',y',z') = g_۱(x,y,z) \quad , \quad G_۲(x,y,z,x',y',z') = g_۲(x',y',z')$$

بدین ترتیب قیدهای مسأله $G_۱(x,y,z,x',y',z') = ۰$ و $G_۲(x,y,z,x',y',z') = ۰$ می‌باشند. توجه کنید که

$$\nabla G_۱(x,y,z,x',y',z') = (\frac{\partial g_۱}{\partial x}, \frac{\partial g_۱}{\partial y}, \frac{\partial g_۱}{\partial z}, ۰, ۰, ۰)$$

$$\nabla G_۲(x,y,z,x',y',z') = (۰, ۰, ۰, \frac{\partial g_۲}{\partial x'}, \frac{\partial g_۲}{\partial y'}, \frac{\partial g_۲}{\partial z'})$$

که مستقل خطی هستند زیرا که $\nabla g_۱ \neq \underline{۰}$ و $\nabla g_۲ \neq \underline{۰}$ در سراسر رویه‌های داده شده فرض شده

است. رابطهٔ لاگرانژ $\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$ به صورت زیر در می‌آید:

$$\begin{cases} 2(x-x') = \lambda_1 \dfrac{\partial g_1}{\partial x} + \lambda_2(0) \\ 2(y-y') = \lambda_1 \dfrac{\partial g_1}{\partial y} + \lambda_2(0) \\ 2(z-z') = \lambda_1 \dfrac{\partial g_1}{\partial z} + \lambda_2(0) \\ -2(x-x') = \lambda_1(0) + \lambda_2 \dfrac{\partial g_2}{\partial x'} \\ -2(y-y') = \lambda_1(0) + \lambda_2 \dfrac{\partial g_2}{\partial y'} \\ -2(z-z') = \lambda_1(0) + \lambda_2 \dfrac{\partial g_2}{\partial z'} \end{cases}$$

اگر $P = (x, y, z)$ و $Q = (x', y', z')$ جواب مسأله باشند، که باید در این روابط صدق کنند، از سه رابطهٔ اول داریم $2\vec{QP} = \lambda_1 \nabla g_1(P)$، پس $\vec{PQ}$ بر $g_1(x, y, z) = 0$ در نقطهٔ P عمود است. همین طور از سه رابطهٔ دوم نتیجه می‌شود که $-2\vec{PQ} = \lambda_2 \nabla g_2(Q)$، یعنی $\vec{PQ}$ در نقطهٔ Q بر $g_2(x', y', z') = 0$ عمود است، و نتیجه مطلوب به‌دست می‌آید.

دو توضیح در مورد مثال بالا بجا است. یکی اینکه حکم مشابه برای اشکال مسطح، مثلاً دو خط متنافر در $\mathbb{R}^3$، را معمولاً در چارچوب هندسه اقلیدسی ثابت می‌کنیم، ولی حتی بیان مطلب برای اشیاء خمیده نیاز به مفهوم "شیء هموار" دارد که در اینجا به تعریفی از آن رسیده‌ایم. نکتهٔ دیگر این که برای حل این مسأله سه بعدی (دو رویه در $\mathbb{R}^3$) عملاً مطلب را به یک مسأله در $\mathbb{R}^6$ تبدیل کردیم و به سهولت به جواب رسیدیم. توجه کنید که $G_1(x, y, z, x', y', z') = 0$ و $G_2(x, y, z, x', y', z') = 0$ هر یک یک "شیء هموار ۵ـ بعدی" در $\mathbb{R}^6$ تعریف می‌کند که اشتراک آنها به سبب استقلال خطی ∇G_1 و ∇G_2 یک "شیء هموار ۴ـ بعدی" در $\mathbb{R}^6$ می‌شود. آنگاه مینیمم تحدید تابع f را به این شیء چهار بعدی مد نظر قرار دادیم.